UTROPHIN GENE PROMOTER

The present invention is based on cloning of a genomic promoter region of the human utrophin gene and of the mouse utrophin gene.

The severe muscle wasting disorders Duchenne muscular dystrophy (DMD) and the less debilitating Becker muscular dystrophy (BMD) are due to mutations in the dystrophin gene resulting in a lack of dystrophin or abnormal expression of truncated forms of dystrophin, respectively. Dystrophin is a large cytoskeletal protein (427kDa with a length of 125nm) which in muscle is located at the cytoplasmic surface of the sarcolemma, the neuromuscular junction (NMJ) and myotendinous junction (MTJ). It binds to a complex of proteins and glycoproteins spanning the sarcolemma called the dystrophin associated glycoprotein complex (DGC). The breakdown of the integrity of this complex due to loss of, or impairment of dystrophin function, leads to muscle degeneration and the DMD phenotype.

The dystrophin gene is the largest gene so far identified in man, covering over 2.7 megabases and containing 79 exons. The corresponding 14kb dystrophin mRNA is expressed predominantly in skeletal, cardiac and smooth muscle with lower levels in brain. Transcription of dystrophin in different tissues is regulated from either the brain promoter (predominantly active in neuronal cells) or muscle promoter (differentiated myogenic cells, and primary glial cells) giving rise to differing first exons. A third promoter between the muscle promoter and the second exon of dystrophin regulates expression in cerebellar Purkinje neurons. Recently reviewed in (Tinsley, *et al* (1994) *Proc Natl Acad Sci U S A* **91**, 8307-13, Blake, *et al* (1994) *Trends in Cell Biol.* **4**: 19-23,Tinsley, *et al* (1993) *Curr Opin Genet Dev.* **3**: 484-90).

There are various approaches which have been adopted for the gene therapy of DMD, using the *mdx* mouse as a model system. However, there are considerable problems related to the number of muscle cells that can be made dystrophin positive, the levels of expression of the gene and the duration of expression (Partridge, *et al.* (1995) *British Medical Bulletin* **51**: 123-137). It has also become apparent that simply re-introducing genes expressing the dystrophin carboxy-terminus has no effect on the dystrophic phenotype although the DGC appears to be re-established at the sarcolemma (Cox, *et al.* (1994) *Nature Genet* **8**: 333-339,Greenberg, *et al.* (1994) *Nature Genet* **8**: 340-344).

In order to circumvent some of these problems, possibilities of compensating for dystrophin loss using a related protein, utrophin, are being explored as an alternative route to dystrophin gene therapy. A similar strategy is currently being evaluated in clinical trials to up-regulate foetal haemoglobin to compensate for the affected adult-globin chains in patients with sickle cell anaemia (Rodgers, *et al.* (1993) *N Engl J Med.* **328**: 73-80,Perrine, *et al.* (1993) *N Engl J Med.* **328**: 81-86).

Utrophin is a 395kDa protein encoded by multiexonic 1Mb *UTRN* gene located on chromosome 6q24 (Pearce, *et al.* (1993) *Hum Mol Gene.* **2**: 1765-1772). At present the tissue regulation of utrophin is not fully understood. In the dystrophin deficient *mdx* mouse, utrophin levels in muscle remain elevated soon after birth compared with normal mice; once the utrophin levels have decreased to the adult levels (about 1 week after birth), the first signs of muscle fibre necrosis are detected. However there is evidence to suggest that in the small calibre muscles, continual increased levels of utrophin can interact with the DGC complex (or an antigenically related complex) at the sarcolemma thus preventing loss of the complex with the result that these muscles appear normal. There is also a

substantial body of evidence demonstrating that utrophin is capable of localising to the sarcolemma in normal muscle. During fetal muscle development there is increased utrophin expression, localised to the sarcolemma, up until 18 weeks in the human and 20 days gestation in the mouse. After this time the utrophin sarcolemmal staining steadily decreases to the significantly lower adult levels shortly before birth where utrophin is localised almost exclusively to the NMJ. The decrease in utrophin expression coincides with increased expression of dystrophin. See reviews (Ibraghimov Beskrovnaya, *et al.* (1992) *Nature* **355**, 696-702.,Blake, *et al.* (1994) *Trends in Cell Biol*,.**4**: 19-23,Tinsley, *et al.* (1993) *Curr Opin Genet Dev.* **3**: 484-90).

Thus, in certain circumstances utrophin can localise to the sarcolemma probably at the same binding sites as dystrophin, through interactions with actin and the DGC. Accordingly, if expression of utrophin is sufficiently elevated, it may maintain the DGC and thus alleviate muscle degeneration in DMD/BMD patients (Tinsley, *et al.* (1993) *Neuromuscul Disord* **3**, 537-9.).

However, manipulation of utrophin expression and screening for molecules able to upregulate expression is hampered by the limited understanding of utrophin expression regulation and its promoters. We have previously isolated a promoter element lying within the CpG island at the 5' end of the utrophin locus that is active in a broad range of cell types and tissues, and shown it to be synaptically regulated *in vivo* (Dennis, et al. (1996) *Nucleic Acids Res* **24**, 1646-52 and WO 96/34101). The sequence contains a consensus N-box, a 6bp motif important in the regulation of other genes expressed at the NMJ (Koike, *et al.* (1995) *Proc Natl Acad Sci USA* **92**, 10624-10628). Localisation of utrophin at the NMJ in mature muscle is partially attributable to enhanced transcription of utrophin at sub-junctional myonuclei, with consequent synaptic

accumulation of mRNA (Gramolini, *et al.* (1997) *J Biol Chem* **272**, 8117-20, Vater, *et al.* (1998) *Molecular and Cellular Neuroscience* **10**, 229-242). The utrophin promoter drives synaptic transcription of a reporter gene in *vivo*; this expression pattern is abolished by point mutations within the N-box (Gramolin, et al. (1998) J Biol Chem **273**, 736-43).

The present inventors hypothesised that utrophin might be transcribed from more than one promoter, an important consideration for the following reasons: First, it may be undesirable to interfere with the mechanisms underlying synaptic regulation of genes, as this might affect expression of other post-synaptic components and impair the structure and function of the NMJ; a promoter without synaptic regulatory elements might be a more suitable target for pharmacological manipulation. Second, cardiac dysfunction is a common feature of the dystrophinopathies (Hoogerwaard, *et al.* (1997) *J Neurol* **244**, 657-63, Sasaki, *et al.* (1998) *Am Heart J* **135**, 937-44); if the cardiac utrophin message was transcribed from a different promoter, then it might prove necessary to up-regulate this. Finally, inclusion of additional regulatory sequences might increase the yield of a screening program to identify small molecules capable of transcriptional activation of utrophin.

We have now identified an alternative promoter lying within the large second intron of the utrophin gene, 50kb 3' to exon 2. The promoter is highly regulated, expressed in a wide range of tissues and has little similarity to the synaptically expressed promoter. This promoter drives transcription of a widely expressed unique first exon that splices into a common full-length mRNA at exon 3. This unique exon (called exon IB) encodes a novel 31 amino acid N-terminus for the utrophin protein which may be involved in binding to the muscle membrane. The sequences of the two utrophin promoters are dissimilar, and we predict that they respond to discrete sets of cellular signals.

Exon IB is primarily considered herein to encode the indicated 31 amino acids. However, the splice occurs within a codon for aspartate. This aspartate residue is common to both isoforms of utrophin. In embodiments of the invention an aspartate residue may be included C-terminal to the 31 amino acids to provide a 32 amino acid peptide, which may be joined to additional amino acids, for instance additional utrophin sequence as discussed. See, for instance, Figure 8 (SEQ ID NO:7) for one embodiment.

These findings significantly contribute to the understanding of the molecular physiology of utrophin expression and are important because the promoter reported here provides an alternative target for transcriptional activation of utrophin in DMD muscle. This promoter does not contain synaptic regulatory elements and might, therefore, be a more suitable target for pharmacological manipulation than the previously described promoter.

We have now cloned this alternative utrophin promoter and exon, and the present invention in various aspects and embodiments is based on the sequence information obtained and provided herein.

One major use of the promoter is in screening for substances able to modulate its activity. It is well known that pharmaceutical research leading to the identification of a new drug generally involves the screening of very large numbers of candidate substances, both before and even after a lead compound has been found. This is one factor which makes pharmaceutical research very expensive and time-consuming. A method or means assisting in the screening process will have considerable commercial importance and utility. Substances identified as upregulators of the utrophin promoter represent an advance in the fight against muscular dystrophy since they provide basis for design and investigation of therapeutics for

in vivo use.

In one aspect, the present invention provides an isolated nucleic acid comprising a promoter, the promoter comprising a sequence of nucleotides shown in Figure 1 (SEQ ID NO:1)or Figure 2 (SEQ ID NO:3). The promoter may comprise one or more fragments of the sequence shown in Figure 1 of Figure 2 sufficient to promote gene expression. The promoter may comprise or consist essentially of a sequence of nucleotides 5' to position 1440 in Figure 1 (human) or position 1183 in Figure 2 (mouse). Preferably the promoter comprises or consists essentially of nucleotides 1199 to 1440 of the human sequence shown in Figure 1, or the equivalent sequence in mouse, e.g. nucleotides 959 to 1183 of Figure 2.

An even smaller portion of this part of the sequences shown in Figure 1 of Figure 2 may be used as long as promoter activity is retained. Restriction enzymes or nucleases may be used to digest the nucleic acid, followed by an appropriate assay (for example as illustrated herein using luciferase constructs) to determine the minimal sequence required. A preferred embodiment of the present invention provides a nucleic acid isolate with the minimal nucleotide sequence shown in Figure 1 or Figure 2 required for promoter activity. The minimal promoter element is situated between the PvuII restriction site at position 1199 in the human sequence and the transcription start site at 1440 bp in the human sequence and between nucleotides 959 to 1183 in the mouse sequence (see Figure 2).

In one embodiment a promoter according to the present invention comprises or consists of sequence that is shown in Figure 3 to be conserved between the human and mouse sequences, e.g. the 25 nucleotide sequence: ACAGGACATCCCAGTGTGCAGTTCG (SEQ ID NO:10) spanning the transcriptional start site.

The promoter may comprise one or more sequence motifs or elements conferring developmental and/or tissue-specific regulatory control of expression. For instance, the promoter may comprise a sequence for muscle-specific expression, e.g. an E-box element/myoD binding site, such as CANNTG, preferably CAGGTG.

Other regulatory sequences may be included, for instance as identified by mutation or digest assay in an appropriate expression system or by sequence comparison with available information, e.g. using a computer to search on-line databases.

By "promoter" is meant a sequence of nucleotides from which transcription may be initiated of DNA operably linked downstream (i.e. in the 3' direction on the sense strand of double-stranded DNA).

"Operably linked" means joined as part of the same nucleic acid molecule, suitably positioned and oriented for transcription to be initiated from the promoter. DNA operably linked to a promoter is "under transcriptional initiation regulation" of the promoter.

The present invention extends to a promoter which has a nucleotide sequence which is allele, mutant, variant or derivative, by way of nucleotide addition, insertion, substitution or deletion of a promoter sequence as provided herein. Systematic or random mutagenesis of nucleic acid to make an alteration to the nucleotide sequence may be performed using any technique known to those skilled in the art. One or more alterations to a promoter sequence according to the present invention may increase or decrease promoter activity, or increase or decrease the magnitude of the effect of a substance able to modulate the promoter activity.

"Promoter activity" is used to refer to ability to initiate transcription. The level of promoter activity is quantifiable for instance by assessment of the amount of mRNA produced by transcription from the promoter or by assessment of the amount of protein product produced by translation of mRNA produced by transcription from the promoter. The amount of a specific mRNA present in an expression system may be determined for example using specific oligonucleotides which are able to hybridise with the mRNA and which are labelled or may be used in a specific amplification reaction such as the polymerase chain reaction. Use of a reporter gene as discussed further below facilitates determination of promoter activity by reference to protein production.

In various embodiments of the present invention a promoter which has a sequence that is a fragment, mutant, allele, derivative or variant, by way of addition, insertion, deletion or substitution of one or more nucleotides, of the sequence of either the human or the mouse promoters shown in Figures 1 and 2, respectively, has at least about 60% homology with one or both of the shown sequences, preferably at least about 70% homology, more preferably at least about 80% homology, more preferably at least about 90% homology, more preferably at least about 95% homology. The sequence in accordance with an embodiment of the invention may hybridise with one or both of the shown sequences, or the complementary sequences (since DNA is generally double-stranded).

Similarity or homology (the terms are used interchangeably) or identity is preferably determined using GAP, from version 20 of GCG. This uses the algorithm of Needleman and Wunsch to align sequences inserting gaps as appropriate to improve the agreement between the two sequences. Parameters employed are the default ones: for nucleotide sequences - Gap Weight 50, Length Weight 3, Average Match 10.000, Average Mismatch 0.000; for peptide sequences - Gap Weight 8, Length Weight 2, Average

Match 2.912, Average Mismatch -2.003. Peptide similarity scores are taken from the BLOSUM62 matrix. Also useful is the TBLASTN program, of Altschul *et al.* (1990) *J. Mol. Biol.* 215: 403-10, or BestFit, which is part of the Wisconsin Package, Version 8, September 1994, (Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA, Wisconsin 53711). Sequence comparisons may be made using FASTA and FASTP (see Pearson & Lipman, 1988. Methods in Enzymology 183: 63-98). Parameters are preferably set, using the default matrix, as follows: Gapopen (penalty for the first residue in a gap): -12 for proteins / -16 for DNA; Gapext (penalty for additional residues in a gap): -2 for proteins / -4 for DNA; KTUP word length: 2 for proteins / 6 for DNA.

Nucleic acid sequence homology may be determined by means of selective hybridisation between molecules under stringent conditions.

Preliminary experiments may be performed by hybridising under low stringency conditions. For probing, preferred conditions are those which are stringent enough for there to be a simple pattern with a small number of hybridisations identified as positive which can be investigated further.

For example, hybridizations may be performed, according to the method of Sambrook et al. (below) using a hybridization solution comprising: 5X SSC (wherein `SSC' = 0.15 M sodium chloride; 0.15 M sodium citrate; pH 7), 5X Denhardt's reagent, 0.5-1.0% SDS, 100 μg/ml denatured, fragmented salmon sperm DNA, 0.05% sodium pyrophosphate and up to 50% formamide. Hybridization is carried out at 37-42°C for at least six hours. Following hybridization, filters are washed as follows: (1) 5 minutes at room temperature in 2X SSC and 1% SDS; (2) 15 minutes at room temperature in 2X SSC and 0.1% SDS; (3) 30 minutes - 1 hour at 37°C in 1X SSC and 1% SDS; (4) 2 hours at 42-65°C in 1X SSC and 1% SDS, changing the solution

every 30 minutes.

One common formula for calculating the stringency conditions required to achieve hybridization between nucleic acid molecules of a specified sequence homology is (Sambrook et al., 1989): T_m = 81.5°C + 16.6Log [Na+] + 0.41 (% G+C) - 0.63 (% formamide) - 600/#bp in duplex.

As an illustration of the above formula, using [Na+] = [0.368] and 50-% formamide, with GC content of 42% and an average probe size of 200 bases, the T_m is 57°C. The T_m of a DNA duplex decreases by 1 - 1.5°C with every 1% decrease in homology. Thus, targets with greater than about 75% sequence identity would be observed using a hybridization temperature of 42°C. Such a sequence would be considered substantially homologous to the nucleic acid sequence of the present invention.

It is well known in the art to increase stringency of hybridisation gradually until only a few positive clones remain. Other suitable conditions include, e.g. for detection of sequences that are about 80-90% identical, hybridization overnight at 42°C in 0.25M Na_2HPO_4, pH 7.2, 6.5% SDS, 10% dextran sulfate and a final wash at 55°C in 0.1X SSC, 0.1% SDS. For detection of sequences that are greater than about 90% identical, suitable conditions include hybridization overnight at 65°C in 0.25M Na_2HPO_4, pH 7.2, 6.5% SDS, 10% dextran sulfate and a final wash at 60°C in 0.1X SSC, 0.1% SDS.

In a further embodiment, hybridisation of nucleic acid molecule to an allele or variant may be determined or identified indirectly, e.g. using a nucleic acid amplification reaction, particularly the polymerase chain reaction (PCR). PCR requires the use of two primers to specifically amplify target nucleic acid, so preferably two nucleic acid molecules

with sequences characteristic of the utrophin promoter are employed. Using RACE PCR, only one such primer may be needed (see "PCR protocols; A Guide to Methods and Applications", Eds. Innis *et al*, *Academic Press, New York*, (1990)).

Thus a method involving use of PCR in obtaining nucleic acid according to the present invention may include:

(a) providing a preparation of nucleic acid, e.g. from a muscle cell;

(b) providing a pair of nucleic acid molecule primers useful in (i.e. suitable for) PCR, at least one of said primers being a primer specific for nucleic acid according to the present invention;

(c) contacting nucleic acid in said preparation with said primers under conditions for performance of PCR;

(d) performing PCR and determining the presence or absence of an amplified PCR product.

The presence of an amplified PCR product may indicate identification of an allele or other variant. The sequence may have the ability to promote transcription (i.e. have "promoter activity") in muscle cells, e.g. human muscle cells, or muscle-specific transcription.

Further provided by the present invention is a nucleic acid construct comprising a utrophin promoter region or a fragment, mutant, allele, derivative or variant thereof able to promoter transcription, operably linked to a heterologous gene, e.g. a coding sequence. By "heterologous" is meant a gene other than utrophin. Modified forms of utrophin are generally excluded. Generally, the gene may be transcribed into mRNA which may be translated into a peptide or polypeptide product which may be detected and preferably quantitated following expression. A gene whose encoded product may be assayed following expression is termed a "reporter gene", i.e. a gene which "reports" on promoter activity.

The reporter gene preferably encodes an enzyme which catalyses a reaction which produces a detectable signal, preferably a visually detectable signal, such as a coloured product. Many examples are known, including β-galactosidase and luciferase. β-galactosidase activity may be assayed by production of blue colour on substrate, the assay being by eye or by use of a spectrophotometer to measure absorbance. Fluorescence, for example that produced as a result of luciferase activity, may be quantitated using a spectrophotometer. Radioactive assays may be used, for instance using chloramphenicol acetyltransferase, which may also be used in non-radioactive assays. The presence and/or amount of gene product resulting from expression from the reporter gene may be determined using a molecule able to bind the product, such as an antibody or fragment thereof. The binding molecule may be labelled directly or indirectly using any standard technique.

Those skilled in the art are well aware of a multitude of possible reporter genes and assay techniques which may be used to determine gene activity. Any suitable reporter/assay may be used and it should be appreciated that no particular choice is essential to or a limitation of the present invention.

Expression of a reporter gene from the promoter may be in an *in vitro* expression system or may be intracellular (*in vivo*). Expression generally requires the presence, in addition to the promoter which initiates transcription, a translational initiation region and transcriptional and translational termination regions. One or more introns may be present in the gene, along with mRNA processing signals (e.g. splice sites).

Systems for cloning and expression of a polypeptide are discussed further below.

The present invention also provides a nucleic acid vector

comprising a promoter as disclosed herein. Such a vector may comprise a suitably positioned restriction site or other means for insertion into the vector of a sequence heterologous to the promoter to be operably linked thereto.

Suitable vectors can be chosen or constructed, containing appropriate regulatory sequences, including promoter sequences, terminator fragments, polyadenylation sequences, enhancer sequences, marker genes and other sequences as appropriate. For further details see, for example, *Molecular Cloning: a Laboratory Manual*: 2nd edition, Sambrook et al, 1989, Cold Spring Harbor Laboratory Press. Procedures for introducing DNA into cells depend on the host used, but are well known.

Thus, a further aspect of the present invention provides a host cell containing a nucleic acid construct comprising a promoter element, as disclosed herein, operably linked to a heterologous gene. A still further aspect provides a method comprising introducing such a construct into a host cell. The introduction may employ any available technique, including, for eukaryotic cells, calcium phosphate transfection, DEAE-Dextran transfection, electroporation, liposome-mediated transfection and transduction using retrovirus.

The introduction may be followed by causing or allowing expression of the heterologous gene under the control of the promoter, e.g. by culturing host cells under conditions for expression of the gene.

In one embodiment, the construct comprising promoter and gene is integrated into the genome (e.g. chromosome) of the host cell. Integration may be promoted by inclusion in the construct of sequences which promote recombination with the genome, in accordance with standard techniques.

Many known techniques and protocols for manipulation of nucleic acid, for example in preparation of nucleic acid constructs, mutagenesis, sequencing, introduction of DNA into cells and gene expression, and analysis of proteins, are described in detail in *Current Protocols in Molecular Biology*, Second Edition, Ausubel et al. eds., John Wiley & Sons, 1994, the disclosure of which is incorporated herein by reference.

Nucleic acid molecules, constructs and vectors according to the present invention may be provided isolated and/or purified (i.e. from their natural environment), in substantially pure or homogeneous form, free or substantially free of a utrophin coding sequence, or free or substantially free of nucleic acid or genes of the species of interest or origin other than the promoter sequence. Nucleic acid according to the present invention may be wholly or partially synthetic. The term "isolate" encompasses all these possibilities.

Nucleic acid constructs comprising a promoter (as disclosed herein) and a heterologous gene (reporter) may be employed in screening for a substance able to modulate utrophin promoter activity. For therapeutic purposes, e.g. for treatment of muscular dystrophy, a substance able to up-regulate expression of the promoter may be sought. A method of screening for ability of a substance to modulate activity of a utrophin promoter may comprise contacting an expression system, such as a host cell, containing a nucleic acid construct as herein disclosed with a test or candidate substance and determining expression of the heterologous gene. The level of transcription of the heterologous gene, or the level of heterologous protein may be determined. The level of protein may be determined by measuring the amount of protein, or the activity of the protein, using techniques known to those skilled in the art.

Alternatively, or additionally a method of screening for

ability of a substance to modulate activity of a utrophin promoter may comprise contacting a cell containing an endogenous utrophin gene (e.g. a mammalian muscle cell) with a test substance and measuring the level of RNA transcription or protein expression using binding members specific for the nucleic acid or polypeptides disclosed herein. Specific binding members include antibodies and nucleic acid probes.

The level of expression in the presence of the test substance may be compared with the level of expression in the absence of the test substance. A difference in expression in the presence of the test substance indicates ability of the substance to modulate gene expression. An increase in expression of the heterologous gene compared with expression of another gene not linked to a promoter as disclosed herein indicates specificity of the substance for modulation of the utrophin promoter.

A promoter construct may be transfected into a cell line using any technique previously described to produce a stable cell line containing the reporter construct integrated into the genome. The cells may be grown and incubated with test compounds for varying times. The cells may be grown in 96 well plates to facilitate the analysis of large numbers of compounds. The cells may then be washed and the reporter gene expression analysed. For some reporters, such as luciferase, the cells will be lysed then analysed. Previous experiments testing the effects of glucocorticoids on the endogenous utrophin protein and RNA levels in myoblasts have already been described [12,13] and techniques used for those experiments may similarly be employed.

Constructs comprising one or more developmental and/or time-specific regulatory motifs (as discussed) may be used to screen for a substance able to modulate the corresponding aspect of the promoter activity, e.g. muscle-specific

expression.

Following identification of a substance which modulates or affects utrophin promoter activity, the substance may be investigated further. Furthermore, it may be manufactured and/or used in preparation, i.e. manufacture or formulation, of a composition such as a medicament, pharmaceutical composition or drug. These may be administered to individuals.

As noted above, the inventors also identified a novel coding sequence (Exon IB) which encodes a novel utrophin N-terminus.

According to a further aspect of the present invention there is provided a nucleic acid molecule which has a nucleotide sequence encoding a polypeptide which includes the amino acid sequence shown in Figure 1 (SEQ ID NO:2) or Figure 2 (SEQ ID NO:4). Such a polypeptide may include other utrophin sequences, and the nucleic acid molecule may be in the form of a utrophin "mini-gene"(discussed further below).

Such a polypeptide may include non-utrophin (i.e. heterologous or foreign) sequences and thereby form a larger fusion protein. For example, such a fusion protein could be used to target a non-utrophin polypeptide to muscle membranes.

The coding sequence included may be that shown in Figure 1 or Figure 2 or it may be a mutant, variant, derivative or allele of the sequence shown. The sequence may differ from that shown by a change which is one or more of addition, insertion, deletion and substitution of one or more nucleotides of the sequence shown. Changes to a nucleotide sequence may result in an amino acid change at the protein level, or not, as determined by the genetic code.

Thus, nucleic acid according to the present invention may

include a sequence different from the sequences shown in Figure 1 or Figure 2 yet encode a polypeptide with the same amino acid sequence. The amino acid sequences shown in Figure 1 and figure 2 consist of 31 residues.

On the other hand the encoded polypeptide may comprise an amino acid sequence which differs by one or more amino acid residues from the amino acid sequences shown in Figure 1 or Figure 2. Nucleic acid encoding a polypeptide which is an amino acid sequence mutant, variant, derivative or allele of the sequences shown in Figure 1 and Figure 2 are further provided by the present invention. Nucleic acid encoding such a polypeptide may show at the nucleotide sequence and/or encoded amino acid level greater than about 60% homology with the coding sequence and/or the amino acid sequence shown in Figure 1 or Figure 2, greater than about 70% homology, greater than about 80% homology, greater than about 90% homology or greater than about 95% homology. Determination of homology is discussed elsewhere herein.

A polypeptide which is a variant, allele, derivative or mutant may have an amino acid sequence which differs from that given in a figure herein by one or more of addition, substitution, deletion and insertion of one or more amino acids. Preferred such polypeptides have wild-type function, that is to say have one or more of the following properties: immunological cross-reactivity with an antibody reactive the polypeptide for which the sequence is given in Figure 1 or Figure 2; sharing an epitope with the polypeptide for which the amino acid sequence is shown in Figure 1 or Figure 2(as determined for example by immunological cross-reactivity between the two polypeptides); a biological activity which is inhibited by an antibody raised against the polypeptide whose sequence is shown in Figure 1 or Figure 2; ability to bind muscle membrane, ability to bind actin; ability to bind DPC.

Variations in amino acid sequence include "conservative

variation", i.e. substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as arginine for lysine, glutamic for aspartic acid, or glutamine for asparagine. Particular amino acid sequence variants may differ from that shown in Figure 1 or Figure 2 by insertion, addition, substitution or deletion of 1 amino acid, 2, 3, 4, or 5-10 amino acids.

According to one aspect of the present invention there is provided a nucleic acid molecule comprising a sequence of nucleotides encoding a polypeptide with utrophin function. Utrophin nucleotide sequences which may be included in the nucleic acid molecule are disclosed in WO 97/922696 which is incorporated herein by reference.

See also Figure 8 and Figure 9 for disclosure of nucleic acid molecules and polypeptides according to the present invention, comprising the exon IB sequence of the invention.

A polypeptide with utrophin function is able to bind actin and able to bind the dystrophin protein complex (DPC).

The nucleic acid molecule may be an isolate, or in an isolated and/or purified form, that is to say not in an environment in which it is found in nature, removed from its natural environment. It may be free from other nucleic acid obtainable from the same species, e.g. encoding another polypeptide.

In one embodiment, nucleic acid molecule is a "mini-gene", i.e. the polypeptide encoded does not correspond to full-length utrophin but is rather shorter, a truncated version (Utrophin mini-genes are discussed in WO97/22696). For instance, part or all of the rod domain may be missing, such

that the polypeptide comprises an actin-binding domain and a DPC-binding domain but is shorter than naturally occurring utrophin. In a full-length utrophin gene including what are identified herein as exons 1A and 1B, the actin-binding domain is encoded by nucleotides 1-739, while the DPC-binding domain (CRCT) is encoded by nucleotides 8499-10301 (where 1 represents the start of translation). See also Figure 8 (SEQ ID NO:5). The respective domains in the polypeptide encoded by a mini-gene according to the invention may comprise amino acids corresponding to those encoded by these nucleotides in the full-length coding sequence. In one embodiment, a minigene according to the present invention comprises or consists of the amino acid sequence encoded by nucleotides 1-739 and 8499-10301 of the A isoform of utrophin in which exon 1B as identified herein is substituted for exons 1A and 2A. The sequence of such a minigene can be constructed by the ordinary skilled person using information disclosed herein, taking into account the content of WO97/22696 and Tinsley et al, *Nature* (1996) 384:349. The nucleic acid sequence and predicted amino acid sequence encoded by a `mini-gene' according to the present invention are shown in Figure 9 (SEQ ID NO:8).

Advantages of a mini-gene over a sequence encoding a full-length utrophin molecule or derivative thereof include easier manipulation and inclusion in vectors, such as adenoviral and retroviral vectors for delivery and expression.

A further preferred non-naturally occurring nucleic acid molecule encoding a polypeptide with the specified characteristics is a chimaeric construct wherein the encoding sequence comprises a sequence obtainable from one mammal, preferably human ("a human sequence"), and a sequence obtainable from another mammal, preferably mouse ("a mouse sequence"). Such a chimaeric construct may of course comprise the addition, insertion, substitution and/or deletion of one or more nucleotides with respect to the parent mammalian

sequences from which it is derived. Preferably, the part of the coding sequence which encodes the actin-binding domain comprises a sequence of nucleotides obtainable from the mouse, or other non-human mammal, or a sequence of nucleotides derived from a sequence obtainable from the mouse, or other non-human mammal.

In a preferred embodiment, the sequence of nucleotides encoding the polypeptide comprises sequence GAGGCAC at residues 331-337 and/or the sequence GATTGTGGATGAAAACAGTGGG (SEQ ID NO:11) at residues 1453-1475 (using the conventional numbering from the initiation codon ATG), and a sequence obtainable from a human.

Nucleic acid according to the present invention is obtainable using one or more oligonucleotide probes or primers designed to hybridise with one or more fragments of a nucleic acid sequence shown in Figure 1 or Figure 2 particularly fragments of relatively rare sequence, based on codon usage or statistical analysis. The amino acid sequence information provided may be used in design of degenerate probes/primers or "long" probes. A primer designed to hybridise with a fragment of the nucleic acid sequence shown may be used in conjunction with one or more oligonucleotides designed to hybridise to a sequence in a cloning vector within which target nucleic acid has been cloned, or in so-called "RACE" (rapid amplification of cDNA ends) in which cDNA's in a library are ligated to an oligonucleotide linker and PCR is performed using a primer which hybridises with the sequence shown in the figures and a primer which hybridises to the oligonucleotide linker.

Nucleic acid isolated and/or purified from one or more cells (e.g. human, mouse) or a nucleic acid library derived from nucleic acid isolated and/or purified from cells (e.g. a cDNA library derived from mRNA isolated from the cells), may be probed under conditions for selective hybridisation and/or subjected to a specific nucleic acid amplification reaction

such as the polymerase chain reaction (PCR).

A method may include hybridisation of one or more (e.g. two) probes or primers to target nucleic acid. Where the nucleic acid is double-stranded DNA, hybridisation will generally be preceded by denaturation to produce single-stranded DNA. The hybridisation may be as part of a PCR procedure, or as part of a probing procedure not involving PCR. An example procedure would be a combination of PCR and low stringency hybridisation. A screening procedure, chosen from the many available to those skilled in the art, is used to identify successful hybridisation events and isolated hybridised nucleic acid.

Probing may employ the standard Southern blotting technique. For instance DNA may be extracted from cells and digested with different restriction enzymes. Restriction fragments may then be separated by electrophoresis on an agarose gel, before denaturation and transfer to a nitrocellulose filter. Labelled probe may be hybridised to the DNA fragments on the filter and binding determined. DNA for probing may be prepared from RNA preparations from cells.

Preliminary experiments may be performed by hybridising under low stringency conditions various probes to Southern blots of DNA digested with restriction enzymes. Suitable conditions would be achieved when a large number of hybridising fragments were obtained while the background hybridisation was low. Using these conditions nucleic acid libraries, e.g. cDNA libraries representative of expressed sequences, may be searched.

It may be necessary for one or more gene fragments to be ligated to generate a full-length coding sequence. Also, where a full-length encoding nucleic acid molecule has not been obtained, a smaller molecule representing part of the

full molecule, may be used to obtain full-length clones. Inserts may be prepared from partial cDNA clones and used to screen cDNA libraries.

Those skilled in the art are well able to employ suitable conditions of the desired stringency for selective hybridisation, taking into account factors such as oligonucleotide length and base composition, temperature and so on. Exemplary conditions have been discussed already above.

Nucleic acid according to the present invention may form part of a cloning vector and/or a vector from which the encoded polypeptide may be expressed. Polypeptide expression is discussed below. Suitable vectors can be chosen or constructed, containing appropriate and appropriately positioned regulatory sequences, as discussed elsewhere herein.

A further aspect of the present invention provides a polypeptide which comprises the amino acid sequence shown in Figure 1 or Figure 2. As mentioned earlier such a polypeptide may include other utrophin sequences or may include heterologous sequences.

Polypeptides which are amino acid sequence variants, alleles, derivatives or mutants are also provided by the present invention. Such polypeptides are discussed elsewhere herein.

The skilled person can use the techniques described herein and others well known in the art to produce large amounts of peptides, for instance by expression from encoding nucleic acid.

In a further aspect the invention provides a method of making a polypeptide, the method including expression from nucleic

acid encoding the polypeptide (generally nucleic acid according to the invention). This may be conveniently be achieved by growing in culture a host cell containing such a vector, under suitable conditions which cause or allow expression of the polypeptide. Polypeptides may also be expressed in in vitro systems such as reticulocyte lysate.

Systems for cloning and expression of a polypeptide in a variety of different host cells are well known. Suitable host cells include bacteria, mammalian cells, yeast and baculovirus systems. Mammalian cell lines available in the art for expression of a heterologous polypeptide include Chinese hamster ovary cells, HeLa cells, baby hamster kidney cells and many others. A common, preferred bacterial host is *E. coli*.

Thus, a further aspect of the present invention provides a host cell containing heterologous nucleic acid encoding a polypeptide as disclosed herein.

The nucleic acid may be integrated into the genome (e.g. chromosome) of the host cell or may be on an extra-chromosomal vector within the cell, or otherwise identifiably heterologous or foreign to the cell.

A still further aspect provides a method comprising introducing such nucleic acid into a host cell. Suitable techniques are discussed elsewhere herein.

The introduction may be followed by causing or allowing expression from the nucleic acid, e.g. by culturing host cells under conditions for expression of the gene.
The polypeptide encoded by the nucleic acid may be expressed from the nucleic acid *in vitro*, e.g. in a cell-free system or in cultured cells, or *in vivo*.
If the polypeptide is expressed coupled to an appropriate signal leader peptide it may be secreted from the cell into

the culture medium.

Peptides can also be generated wholly or partly by chemical synthesis. The compounds of the present invention can be readily prepared according to well-established, standard liquid or, preferably, solid-phase peptide synthesis methods, general descriptions of which are broadly available (see, for example, in J.M. Stewart and J.D. Young, Solid Phase Peptide Synthesis, 2nd edition, Pierce Chemical Company, Rockford, Illinois (1984), in M. Bodanzsky and A. Bodanzsky, The Practice of Peptide Synthesis, Springer Verlag, New York (1984); and Applied Biosystems 430A Users Manual, ABI Inc., Foster City, California), or they may be prepared in solution, by the liquid phase method or by any combination of solid-phase, liquid phase and solution chemistry, e.g. by first completing the respective peptide portion and then, if desired and appropriate, after removal of any protecting groups being present, by introduction of the residue X by reaction of the respective carbonic or sulfonic acid or a reactive derivative thereof.

The present invention also includes active portions, fragments, derivatives and functional mimetics of the polypeptides of the invention. An "active portion" of a polypeptide means a peptide which is less than said full length polypeptide, but which retains a biological activity, such as a biological activity selected from binding to ligand, binding to muscle membrane. Such an active fragment may be included as part of a fusion protein, e.g. including a polypeptide which is to be targetted to the muscle membrane.

A "fragment" of a polypeptide generally means a stretch of amino acid residues of about five to twenty-five contiguous amino acids, typically about ten to twenty contiguous amino acids. Fragments of the novel N-terminus polypeptide sequence may include antigenic determinants or epitopes useful for

raising antibodies to a portion of the amino acid sequence, or may be sequence useful for targetting to muscle membrane. Alanine scans are commonly used to find and refine peptide motifs within polypeptides, this involving the systematic replacement of each residue in turn with the amino acid alanine, followed by an assessment of biological activity.

Preferred fragments of exon IB polypeptide include those comprising or consisting of an epitope which may be used for instance in raising or isolating antibodies. Variant and derivative peptides, peptides which have an amino acid sequence which differs from one of these sequences by way of addition, insertion, deletion or substitution of one or more amino acids are also provided by the present invention.

A "derivative" of a polypeptide or a fragment thereof may include a polypeptide modified by varying the amino acid sequence of the protein, e.g. by manipulation of the nucleic acid encoding the protein or by altering the protein itself. Such derivatives of the natural amino acid sequence may involve one or more of insertion, addition, deletion or substitution of one or more amino acids, which may be without fundamentally altering the qualitative nature of biological activity of the wild type polypeptide. Also encompassed within the scope of the present invention are functional mimetics of active fragments of the exon 1B polypeptides provided (including alleles, mutants, derivatives and variants). The term "functional mimetic" means a substance which may not contain an active portion of the relevant amino acid sequence, and probably is not a peptide at all, but which retains in qualitative terms biological activity of natural exon 1B polypeptide. The design and screening of candidate mimetics is described in detail below.

A polypeptide according to the present invention may be isolated and/or purified (e.g. using an antibody) for instance

after production by expression from encoding nucleic acid (for which see below). Thus, a polypeptide may be provided free or substantially free from contaminants with which it is naturally associated (if it is a naturally-occurring polypeptide). A polypeptide may be provided free or substantially free of other polypeptides. Polypeptides according to the present invention may be generated wholly or partly by chemical synthesis. The isolated and/or purified polypeptide may be used in formulation of a composition, which may include at least one additional component, for example a pharmaceutical composition including a pharmaceutically acceptable excipient, vehicle or carrier. A composition including a polypeptide according to the invention may be used in prophylactic and/or therapeutic treatment as discussed below.

A polypeptide, peptide, allele, mutant, derivative or variant according to the present invention may be used as an immunogen or otherwise in obtaining specific antibodies. Antibodies are useful in purification and other manipulation of polypeptides and peptides, diagnostic screening and therapeutic contexts.

Accordingly, a further aspect of the present invention provides an antibody able to bind specifically to the polypeptide whose sequence is given in Figure 1 or Figure 2. Such an antibody may be specific in the sense of being able to distinguish between the polypeptide it is able to bind and other human (or mouse) polypeptides for which it has no or substantially no binding affinity (e.g. a binding affinity of about 1000x less). Specific antibodies bind an epitope on the molecule which is either not present or is not accessible on other molecules. Antibodies according to the present invention may be specific for the wild-type polypeptide. Antibodies according to the invention may be specific for a particular mutant, variant, allele or derivative polypeptide as between that molecule and the wild-type polypeptide, so as

to be useful in diagnostic and prognostic methods as discussed below. Antibodies are also useful in purifying the polypeptide or polypeptides to which they bind, e.g. following production by recombinant expression from encoding nucleic acid.

Preferred antibodies according to the invention are isolated, in the sense of being free from contaminants such as antibodies able to bind other polypeptides and/or free of serum components. Monoclonal antibodies are preferred for some purposes, though polyclonal antibodies are within the scope of the present invention.

Antibodies may be obtained using techniques which are standard in the art. Methods of producing antibodies include immunising a mammal (e.g. mouse, rat, rabbit, horse, goat, sheep or monkey) with the protein or a fragment thereof. Antibodies may be obtained from immunised animals using any of a variety of techniques known in the art, and screened, preferably using binding of antibody to antigen of interest. For instance, Western blotting techniques or immunoprecipitation may be used (Armitage et al., 1992, Nature 357: 80-82). Isolation of antibodies and/or antibody-producing cells from an animal may be accompanied by a step of sacrificing the animal.

As an alternative or supplement to immunising a mammal with a peptide, an antibody specific for a protein may be obtained from a recombinantly produced library of expressed immunoglobulin variable domains, e.g. using lambda bacteriophage or filamentous bacteriophage which display functional immunoglobulin binding domains on their surfaces; for instance see WO92/01047. The library may be naive, that is constructed from sequences obtained from an organism which has not been immunised with any of the proteins (or fragments), or may be one constructed using sequences obtained

from an organism which has been exposed to the antigen of interest.

Antibodies according to the present invention may be modified in a number of ways. Indeed the term "antibody" should be construed as covering any binding substance having a binding domain with the required specificity. Thus the invention covers antibody fragments, derivatives, functional equivalents and homologues of antibodies, including synthetic molecules and molecules whose shape mimicks that of an antibody enabling it to bind an antigen or epitope.

Example antibody fragments, capable of binding an antigen or other binding partner are the Fab fragment consisting of the VL, VH, Cl and CH1 domains; the Fd fragment consisting of the VH and CH1 domains; the Fv fragment consisting of the VL and VH domains of a single arm of an antibody; the dAb fragment which consists of a VH domain; isolated CDR regions and F(ab')2 fragments, a bivalent fragment including two Fab fragments linked by a disulphide bridge at the hinge region. Single chain Fv fragments are also included.

A hybridoma producing a monoclonal antibody according to the present invention may be subject to genetic mutation or other changes. It will further be understood by those skilled in the art that a monoclonal antibody can be subjected to the techniques of recombinant DNA technology to produce other antibodies or chimeric molecules which retain the specificity of the original antibody. Such techniques may involve introducing DNA encoding the immunoglobulin variable region, or the complementarity determining regions (CDRs), of an antibody to the constant regions, or constant regions plus framework regions, of a different immunoglobulin. See, for instance, EP184187A, GB 2188638A or EP-A-0239400. Cloning and expression of chimeric antibodies are described in EP-A-0120694 and EP-A-0125023.

Hybridomas capable of producing antibody with desired binding characteristics are within the scope of the present invention, as are host cells, eukaryotic or prokaryotic, containing nucleic acid encoding antibodies (including antibody fragments) and capable of their expression. The invention also provides methods of production of the antibodies including growing a cell capable of producing the antibody under conditions in which the antibody is produced, and preferably secreted.

The reactivities of antibodies on a sample may be determined by any appropriate means. Tagging with individual reporter molecules is one possibility. The reporter molecules may directly or indirectly generate detectable, and preferably measurable, signals. The linkage of reporter molecules may be directly or indirectly, covalently, e.g. via a peptide bond or non-covalently. Linkage via a peptide bond may be as a result of recombinant expression of a gene fusion encoding antibody and reporter molecule.

One favoured mode is by covalent linkage of each antibody with an individual fluorochrome, phosphor or laser dye with spectrally isolated absorption or emission characteristics. Suitable fluorochromes include fluorescein, rhodamine, phycoerythrin and Texas Red. Suitable chromogenic dyes include diaminobenzidine.

Other reporters include macromolecular colloidal particles or particulate material such as latex beads that are coloured, magnetic or paramagnetic, and biologically or chemically active agents that can directly or indirectly cause detectable signals to be visually observed, electronically detected or otherwise recorded. These molecules may be enzymes which catalyse reactions that develop or change colours or cause changes in electrical properties, for example. They may be molecularly excitable, such that electronic transitions

between energy states result in characteristic spectral absorptions or emissions. They may include chemical entities used in conjunction with biosensors. Biotin/avidin or biotin/streptavidin and alkaline phosphatase detection systems may be employed.

The mode of determining binding is not a feature of the present invention and those skilled in the art are able to choose a suitable mode according to their preference and general knowledge. Particular embodiments of antibodies according to the present invention include antibodies able to bind and/or which bind specifically, e.g. with an affinity of at least 10^{-7} M, to the peptides shown in Figure 1 (SEQ ID NO:2) or Figure 2 (SEQ ID NO:4).

Antibodies according to the present invention may be used in screening for the presence of a polypeptide, for example in a test sample containing cells or cell lysate as discussed, and may be used in purifying and/or isolating a polypeptide according to the present invention, for instance following production of the polypeptide by expression from encoding nucleic acid therefor.

An antibody may be provided in a kit, which may include instructions for use of the antibody, e.g. in determining the presence of a particular substance in a test sample. One or more other reagents may be included, such as labelling molecules, buffer solutions, elutants and so on. Reagents may be provided within containers which protect them from the external environment, such as a sealed vial.

The present invention extends in various aspects not only to a substance identified using a nucleic acid molecule as a modulator of utrophin promoter activity, or to a polypeptide, or nucleic acid molecule in accordance with what is disclosed herein, but also a pharmaceutical composition, medicament, drug or other composition comprising such a substance, a

method comprising administration of such a composition to a patient, e.g. for increasing utrophin expression for instance in treatment of muscular dystrophy, use of such a substance in manufacture of a composition for administration, e.g. for increasing utrophin expression for instance in treatment of muscular dystrophy, and a method of making a pharmaceutical composition comprising admixing such a substance with a pharmaceutically acceptable excipient, vehicle or carrier, and optionally other ingredients.

Administration will preferably be in a "therapeutically effective amount", this being sufficient to show benefit to a patient. Such benefit may be at least amelioration of at least one symptom. The actual amount administered, and rate and time-course of administration, will depend on the nature and severity of what is being treated. Prescription of treatment, eg decisions on dosage etc, is within the responsibility of general practitioners and other medical doctors.

A composition may be administered alone or in combination with other treatments, either simultaneously or sequentially dependent upon the condition to be treated.

Pharmaceutical compositions according to the present invention, and for use in accordance with the present invention, may comprise, in addition to active ingredient, a pharmaceutically acceptable excipient, carrier, buffer, stabiliser or other materials well known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active ingredient. The precise nature of the carrier or other material will depend on the route of administration, which may be oral, or by injection, e.g. cutaneous, subcutaneous or intravenous.

Pharmaceutical compositions for oral administration may be in

tablet, capsule, powder or liquid form. A tablet may comprise a solid carrier such as gelatin or an adjuvant. Liquid pharmaceutical compositions generally comprise a liquid carrier such as water, petroleum, animal or vegetable oils, mineral oil or synthetic oil. Physiological saline solution, dextrose or other saccharide solution or glycols such as ethylene glycol, propylene glycol or polyethylene glycol may be included.

For intravenous, cutaneous or subcutaneous injection, or injection at the site of affliction, the active ingredient will be in the form of a parenterally acceptable aqueous solution which is pyrogen-free and has suitable pH, isotonicity and stability. Those of relevant skill in the art are well able to prepare suitable solutions using, for example, isotonic vehicles such as Sodium Chloride Injection, Ringer's Injection, Lactated Ringer's Injection. Preservatives, stabilisers, buffers, antioxidants and/or other additives may be included, as required.

Instead of a substance identified using a promoter as disclosed herein, a mimetic or mimick or the substance may be designed for pharmaceutical use. The designing of mimetics to a known pharmaceutically active compound is a known approach to the development of pharmaceuticals based on a "lead" compound. This might be desirable where the active compound is difficult or expensive to synthesise or where it is unsuitable for a particular method of administration, eg peptides are unsuitable active agents for oral compositions as they tend to be quickly degraded by proteases in the alimentary canal. Mimetic design, synthesis and testing may be used to avoid randomly screening large number of molecules for a target property.

There are several steps commonly taken in the design of a mimetic from a compound having a given target property.

Firstly, the particular parts of the compound that are critical and/or important in determining the target property are determined. In the case of a peptide, this can be done by systematically varying the amino acid residues in the peptide, eg by substituting each residue in turn. These parts or residues constituting the active region of the compound are known as its "pharmacophore".

Once the pharmacophore has been found, its structure is modelled to according its physical properties, eg stereochemistry, bonding, size and/or charge, using data from a range of sources, eg spectroscopic techniques, X-ray diffraction data and NMR. Computational analysis, similarity mapping (which models the charge and/or volume of a pharmacophore, rather than the bonding between atoms) and other techniques can be used in this modelling process. In a variant of this approach, the three-dimensional structure of the ligand and its binding partner are modelled. This can be especially useful where the ligand and/or binding partner change conformation on binding, allowing the model to take account of this the design of the mimetic.

A template molecule is then selected onto which chemical groups which mimic the pharmacophore can be grafted. The template molecule and the chemical groups grafted on to it can conveniently be selected so that the mimetic is easy to synthesise, is likely to be pharmacologically acceptable, and does not degrade *in vivo*, while retaining the biological activity of the lead compound. The mimetic or mimetics found by this approach can then be screened to see whether they have the target property, or to what extent they exhibit it. Further optimisation or modification can then be carried out to arrive at one or more final mimetics for *in vivo* or clinical testing.

Mimetics of substances identified as having ability to

modulate utrophin promoter activity using a screening method as disclosed herein are included within the scope of the present invention.

Modifications to and further aspects and embodiments of the present invention will be apparent to those skilled in the art. All documents mentioned herein are incorporated by reference.

Experimental basis for and embodiments of the present invention will now be described in more detail, by way of example and not limitation, and with reference to the following figures:

Figure 1 shows the sequence of the human exon 1B and promoter B. Numbering corresponds to the insert of pBSX2.0. The deduced translation of exon 1B is shown. The positions of features such as restriction sites, IL-6 response element and Alu repetitive elements are shown.

Figure 2 shows the sequence of the mouse exon 1B and promoter B. Numbering corresponds to the insert of pBSX8.0. The deduced translation of exon 1B is shown. The positions of features such as restriction sites, IL-6 response element and Alu repetitive elements are shown.

Figure 3 shows the sequence alignment of human (top) and mouse (bottom) exon 1B (in upper case) and promoter B. Numbering corresponds to the inserts of pBSX2.0 and pBSX8.0, respectively. The human PvuII site (see Figure 7) is indicated. The open triangle indicates the position at which the luciferase coding sequence was inserted to make pGL3/UtroB/F (see below). The deduced translation of exon 1B is shown; amino acids marked in bold type are identical between the human and mouse sequences. The conserved splice donor consensus is shown in grey. Two putative Ap1 sites and

an initiator-like element (Inr) are 100% conserved and indicated in black. A solid arrow marks the single transcription start indicated by primer extension; figures adjacent to the sequence indicate the number of individual 5'RACE clones that terminated at the positions shown.

Figure 4 shows the position of the primers used in RT-PCR of exon 1B-containing utrophin transcript, and the probes used to probe the PCR products. Primers specific to exon 1B (BF31) and utrophin C-terminus (CT2) were used to amplify 9816bp of utrophin cDNA. The products were blotted and probed with U41, U107, BR4 and U16 as indicated. The diagram is not to scale; numbering refers to the nucleotide sequence of the full-length cDNA. The corresponding functional domains of the protein are indicated above: actin binding domain; rod, rod domain; Cys, cysteine rich domain, C-Term; C-terminal domain.

Figure 5 shows a schematic representation of (A) human YAC and (B) mouse PAC contigs showing position of exons within the genomic map. Key to mouse restriction sites: C, ClaI; S, SacII; B, BssHII; X, XhoI. (C) shows the nomenclature for utrophin promoters, exons and transcripts.

Figure 6 shows the *in vitro* activity of utrophin promoter B. (A) shows normalised luciferase activity following transfection of three different human cell types with either pGL3/utroB/F (`forward construct') or pGL3/utroB/R (`reverse construct').

Figure 7 shows deletion analysis of promoter B. The 1.5kb insert of pGL3/utroB/F was deleted at its 5' and 3' ends using the internal restriction sites indicated. Reporter activity was assayed following transient transfection of IN157 and CL11T47 cells.

Figure 8 shows conceptual translation of exon 1B as part of

utrophin, showing a nucleotide sequence and encoded polypeptide according to embodiments of the present invention.

Figure 9 shows the nucleic acid and predicted amino acid sequence of a utrophin B isoform `minigene'.

Figure 10 shows the dosage dependence of IL-6 mediated expression from the isoform B promoter.

Oligonucleotides, PCR, RT-PCR and 5'RACE

PCR and RT-PCR were performed as described (Blake, *et al.* (1996) *J Biol Chem* **271**, 7802-7810). Oligonucleotide sequences (5' to 3') were:

UM83	gatgttcctg tgaggccttc gag,	(SEQ ID NO:12)
UM82	cactcttgga aaatcgagcg t,	(SEQ ID NO:13)
U16	actatgatgt ctgccagagt tg,	(SEQ ID NO:14)
U107	gatccaatag cttccttcca tcttt,	(SEQ ID NO:15)
UBF	tggaaaaagt ggaggttgga,	(SEQ ID NO:16)
BR2	tccaacctcc actttttcca,	(SEQ ID NO:17)
BR4	gcctggagag ctacatgccc t,	(SEQ ID NO:18)
BF8	ctccacatct ttttcctcat catch,	(SEQ ID NO:19)
BF9	gattgtggtg atggttgtag aa,	(SEQ ID NO:20)
BR10	gattgtggtg atggttgtag aa,	(SEQ ID NO:20)
BR14	gatgatgagg aaaaagatgt ggag,	(SEQ ID NO:21)
BF15	aaacccaaaa taacacagga catc,	(SEQ ID NO:22)
BF16	agtgtaactt ctctctggtg,	(SEQ ID NO:23)
BF31	taagcagatg taggtgatga gc,	(SEQ ID NO:24)
BF42	gctgcttttg ttgtccactt c,	(SEQ ID NO:25)
BR43	atagcttcct tccatctttg ag,	(SEQ ID NO:26)
CT2	ctccacgttc ttccctctct act,	(SEQ ID NO:27)
2ApF	gcgtgcagtg gaccattttt cagattta,	(SEQ ID NO:28)
1BpF	cgctgcagca gccaccacat ttcgttg,	(SEQ ID NO:29)
3pR	gcgtgcagat cgagcgttta tccatttg.	(SEQ ID NO:30)

5' RACE was undertaken using adapter-ligated mouse heart cDNA (Marathon-Ready, Clontech), following the manufacturer's

protocol, using the supplied adapter primers with nested mouse utrophin primers UM83 (exon 4) and UM82 (exon 3). Products were cloned in pGEM-T (Promega). Human exon 1B was isolated from skeletal muscle cDNA by PCR using mouse primers UBF and UM83. 5'RACE was used to clone the 5' end of human exon 1B, using primers U107 and BR4. Full-length utrophin RT-PCR was done as described (Blake, *et al.* (1996) *J Biol Chem* **271**, 7802-7810.), but using Boehringer Expand Reverse Transcriptase and Long Template PCR reagents, and a primer annealing temperature of 59°C. Semi-quantitative RT-PCR was performed using primers BF42 and BR43 to amplify utrophin B, and commercial primers (Stratagene) to amplify glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Exponential amplification was established by withdrawing samples from thermal cycling at 1 cycle intervals over a range of 5 cycles, predicted to span the exponential range following initial experiments in which samples were withdrawn at 5 cycle intervals. Products were blotted and probed with labelled BR4 or a 600bp GA3PH probe. Band intensities were quantified using a Storm phosphoimager. A graph of $\log_2$[band intensity] versus cycle number showed a linear relationship with gradient $\pm$ 1, indicating near-perfect exponential amplification. The band intensities at any given cycle over this range are therefore directly proportional to the amount of cDNA in the original samples.

Genomic Mapping and Clones

Human YACs are as previously described (Pearce, *et al.* (1993) *Hum Mol Genet* **2**, 1765-72). Southern blots of restriction digested YAC DNA were probed with end-labelled BR4. A 3.0kb hybridising XbaI fragment was cloned from YAC 4X124H10 (a YAC clone which contains a human genomic DNA insert)into pBlueScript (Stratagene) generating pBSX2.0. Mouse PACs were identified from the RPCI21 library. A 398bp exon 1B/promoter B DNA probe (UB400) encompassing human positions 1129 to 1527 was used for exon 1B mapping. Library filters were screened with probes to exons 1A-5 (Dennis, *et al.* (1996) *Nucleic Acid*

Res **24**, 1646-52) and UB400. Eleven PACs were identified, and four of these arranged into a contig by restriction mapping. An 8.0kb XbaI fragment from PAC 110C24, that hybridised with UB400, was cloned in pBlueScript generating pBSX8.0.

Northern Blots and Probes

A human multiple tissue northern blot and b-actin control cDNA probe were obtained from Clontech. A utrophin C-terminal cDNA probe, encompassing the last 4.0kb of the utrophin message, was generated by PCR. Human exon 1B sequence between positions 1480 and 1596 was cloned into pGEM-T and an exon 1B antisense riboprobe was transcribed (In Vitro Transcription Kit, Promega) from the SP6 promoter following linearisation of the plasmid with NcoI. Hybridisation was carried out at 70°C in 50% formamide hybridisation buffer (Ausubel, *et al.* (1999) *Current Protocols in Molecular Biology* (Wiley).) and the filter was washed at 75°C in 0.1xSSC, 0.1%SDS for 2 hours.

RNase Protection

Specific probes spanning the exon 1B/3 and exon 2A/3 boundaries were obtained by PCR amplification of mouse heart cDNA using primers 2ApF, 1BpF and 3pR. Products were cloned in the PstI site of pDP18 (Ambion) and sequenced. Plasmids were linearised with EcoR1 (1B) or BamH1 (2A); labelled antisense riboprobe was transcribed from the T7 promoter and gel purified. RNase protection was carried out using RPAIII kit (Ambion) following the manufacturer's instructions (30μg total RNA unless stated, hybridisation temperature 42°C, RNase A/T1 dilution 1:200). Following electrophoretic separation, band intensities were quantified as above, and corrected for the amount of label present in each protected fragment.

Promoter/Reporter Constructs

Reporter constructs were generated by PCR amplification of the human sequence between positions 39 and 1503, using pBSX2.0 as template. Pfu polymerase was used with primers BF9 and BR14.

Following 15 cycles of 96°C for 45 seconds, 62°C for 45 seconds, 72°C for 4 minutes, products were dA-tailed and cloned in pGEM-T. Clones were identified with product in both orientations and insert, liberated by digestion with SacI/NcoI, was cloned into the SacI/NcoI sites of a promoterless luciferase reporter plasmid (pGL3 basic, Promega), generating constructs with insert in forward (pGL3/utroB/F) and reverse (pGL3/UtroB/R) orientation with respect to the coding sequence of luciferase. Deletions of the forward construct were generated by cleavage at SpeI, NdeI, EcoRI and PvuII sites in the insert, followed by religation to sites in the 5' or 3' polylinker. Constructs were sequenced completely.

Cell Culture and Transfections

Three human cell lines (IN157 rhabdomyosarcoma (Nielsen et al., 1993, *Mol Cell Endocrinol* 93: 87-95), CL11T47 kidney epithelial and HeLa cervical epithelial (*Cancer Research, 1952* 12: 264)were maintained as described (Dennis, *et al.* (1996) *Nucleic Acid Res* **24**, 1646-52). 2μg pGL3/utroB/F or R, or its molar equivalent, mixed with 0.5μg of LacZ control plasmid (pSV-β-gal, Promega) was transfected in each well of 6 well plates using Superfect (Qiagen), following the manufacturer's protocol. 48 hours later, cells were harvested and cell extracts were assayed for luciferase and β-galactosidase activity as described (Dennis, *et al.* (1996) *Nucleic Acids Res* **24**, 1646-52). Luciferase activity was standardised to β-galactosidase activity in each individual sample to control for transfection efficiency. Results are expressed as mean luciferase/β-galactosidase ratio for four individual transfections. Error bars indicate the standard error of the mean. For comparison of different constructs within the same cell line, results were standardised to those obtained with pGL3/utroB/F and are expressed as % of this value. For comparison of constructs between cell lines, results were standardised to those obtained with a luciferase-SV40

promoter/enhancer plasmid (pGL3 control, Promega) that generates high levels of reporter activity in all cell lines tested.

Primer Extension

Primer extension was carried out as described (18); end-labelled primer BR2 was annealed to 0, 30 or 50μg mouse heart total RNA at 58°C for 20 minutes, and extended at 42°C for 40 minutes. Products were separated on a 6% polyacrylamide gel, under denaturing conditions, alongside a sequencing ladder generated from pBSX8.0 using primer BR2.

Results

An alternative 5' exon in utrophin mRNA

Utrophin from a mouse heart cDNA library was amplified by 5'RACE, and the resulting products cloned and sequenced. Of 12 clones, 8 contained novel sequence 5' of exon 3. Below, we present evidence that the novel sequence is a single alternative 5' exon of utrophin containing a translational initiation codon. We refer to this sequence as `exon 1B' to distinguish it from the previously described 5' cDNA sequence comprising untranslated exon 1A and exon 2A which contains the translational start (Figure 5c).

Figure 3 shows a sequence comparison of human and mouse exon 1B, and genomic flanking sequence. The position and phase of the splice junction at the 5' end of exon 3 is identical for both exon 1B- and exon 2A- containing transcripts. Exon 1B contains a putative ATG translation initiation codon and open reading frame, in-frame with that of exon 3, predicting a novel 31 amino acid N-terminus to the utrophin protein. The context of the ATG codon is predicted to be favourable for translation in that there is a purine at position -3 (bold in Figure. 3) (33). Human and mouse exons 1B show 82% nucleotide identity. The predicted translations are 84% identical and 94%

similar. The position and context of the ATG codon are conserved. The human sequence contains a second putative ATG codon immediately 5' (position 1511, solid bar in Figure.1), followed by a TAG stop codon. As this ATG does not adhere to the Kozak consensus, is not associated with an open reading frame and is not present in the mouse sequence, we predict that this is not a functional translation start. A similar feature is present in human exon 2A, where the 5'UTR contains a short open reading frame prior to the true translation start.

The transcript associated with exon 1B

A human multiple tissue northern blot was probed with an exon 1B anti-sense riboprobe. A single hybridising 13kb band was observed, identical to that produced by probing the same blot with a cDNA encompassing 4kb of the utrophin C-terminus, indicating that exon1B is exclusively associated with a full-length utrophin mRNA. Exon 1B is ubiquitously expressed, and appears most abundant in heart and pancreas, and least abundant in the brain, relative to β-actin. This is similar to the expression profile of total full-length utrophin.

RT-PCR was employed to confirm the association of exon 1B with a utrophin mRNA predicted to give rise to functional protein (Figure.4). Amplification of first strand cDNA from IN157 cells utilising a forward primer specific to exon 1B (BF3I)and a reverse primer within the utrophin C-terminus (CT2)produced a product of expected size. Successive hybridisation of this PCR product with domain-specific probes; U41, UBR4, U107 and U16, confirmed that exon 1B is associated with a utrophin transcript spanning the full coding sequence of the gene.

The expression profiles of exons 1B and 2A were examined using RNase protection. Specific riboprobes corresponding to the exon 1B/3 and 2A/3 boundaries were simultaneously hybridised with total RNA, allowing direct quantitation of transcript

abundance. B-utrophin is the most abundant form in the heart, whereas exon 2A-containing transcripts predominate in the kidney. Approximately equal amounts of exons 1B and 2A were observed in the brain and in skeletal muscle.

Mapping and cloning of genomic sequence associated with exon 1B

Using probe BR4, exon 1B was mapped within our previously described human YAC contig (26) encompassing the 5' end of the utrophin locus (Figure.5a). A hybridising band was seen with YAC 4X124H10 but not 4X23E3 or 5C2 indicating that exon 1B lies within the 120kb intron 2 of the utrophin gene. A subsequent database search identified a clone from the HGMP human chromosome 6 sequencing project, containing exons 1A, 2A and 1B. This indicated that exon 1B lies 52.2kb 3' of exon 2A (Figure.5a). Probing the mouse genomic PAC library (RPCI21 from P. DeJong, Roswell Park Cancer Institute) with utrophin exons 1A, 1B and 2- 5 inclusive identified a series of genomic PACs spanning the 5' end of the mouse utrophin gene. Four of these PACs were assembled into a contig of the region. Hybridisation with UB400 confirmed that exon 1B lies within intron 2 in the mouse (Figure.5b), approximately 50kb 3' of exon 2.

Human and mouse genomic fragments were obtained from the YAC and PAC libraries, respectively. Genomic sequence encompassing exon 1B was obtained by an Xba I digest of YAC 4X124H10 (human 3kb fragment)and PAC110c24 (mouse 8.8kb fragment). These fragments were sub-cloned into pBluescript vector, the human fragment was deleted to 2kb during the sub-cloning. The plasmid clones were designated pBSX2.0 (human) and pBSX8.0 (mouse). Comparison of the cDNA and genomic sequence showed no evidence of a further 5' exon in the transcript associated with exon 1B, suggesting that the genomic flanking sequence contained the transcription start and promoter element responsible for exon 1B expression. Our

nomenclature for utrophin 5' exons, transcripts and promoters appears in Figure 5c.

Promoter B

1.5kb of human genomic sequence 5' of exon 1B, including the 5'UTR of exon 1B, was cloned in both orientations into a promoterless luciferase reporter vector. Three human cell lines (IN157 rhabdomyosarcoma, CL11T47 kidney epithelial and HeLa cervical epithelial) were transiently transfected with these constructs. These three lines were chosen because they are known to express utrophin mRNA and protein at different levels. Reporter activity was detected at significantly higher levels in cells transfected with the forward than the reverse orientation construct, indicating promoter activity (Figure 6). Interestingly, the level of activity varied between cell lines by an order of magnitude. Semi-quantitative RT-PCR demonstrated that the variation of luciferase expression mimicked the transcription profile of endogenous utrophin exon 1B. In contrast, the GA3PDH control showed identical amplification in all cDNA samples, indicating that the differences seen in B-utrophin amplification have arisen from differences in the level of expression of the endogenous B-utrophin transcript in these cells lines. These data show that the 1.5kb of genomic sequence 5' of exon 1B utilised in these reporter clones contains the necessary signals to initiate transcription of exon 1B, and regulatory elements that determine the level of expression in these cell lines.

To further delineate important elements within this region, a series of 5' and 3' deletions of promoter B were made, and the *in vitro* activity of each one assayed (Figure 7). A 300bp element, contained within clone pGL3/utroB/F/D5' Pvu 1199, retains 70% activity of the full 1.5kb construct in expressing cell lines, and shows 74% identity between human and mouse (Figure.3). Homology falls to 50% when sequence further 5' if the human PvuII site is compared with corresponding mouse

sequence using a 35bp window. Homology was determined using GAP, from version 20 of GCG, with default parameters as noted already above.

Promoter B transcription start site

The 5' ends of 8 human and 4 mouse 5'RACE clones clustered around a putative cap site in the genomic sequence (Figure.3). None of the 5'RACE clones generated by amplification across the exon 3/exon 1B boundary extended further upstream. RT-PCR was carried out using forward primers around this region with a reverse primer in exon 4. A product of expected size was amplified from IN157 cDNA by primers BF42 and BF8, but not BF16 or BF15, indicating that the transcription start is within the 18bp that separates the two primers BF15 and BF42. These 18 bases contain the putative cap site and the cluster of RACE clone 5' ends.

To map the start site accurately, primer extension using an exon 1B reverse primer and mouse heart RNA was employed. This yielded a single product, indicative of a single transcription start site. Transcription initiates at mouse position 1183 within a 25-bp motif, which is 100% conserved between human and mouse. Part of this motif, spanning the cap site, is a 6/7 base match for the initiator consensus, and correspondingly shows homology to the initiators of other genes. The transcription start site is homologous to the initiators of other promoters. Consensus 1, initiator consensus derived from sequence comparison of Inr^+ genes (Azizkhan, et al. (1993) *Critical Reviews in Eukaryotic Gene Expression* **3**, 229-254.); consensus 2, experimentally-derived consensus for functional initiator (Javahery, *et al.* (1994) *Molecular and Cellular Biology* **14**, 116-127.); TdT, terminal deoxynucleotidyl transferase; hRAR, human retinoic acid receptor a; mCREB, mouse cAMP response element binding protein. Transcribed sequence is indicated in bold uppercase. We consider this promoter to be of the $TATA^-Inr^+$ type.

Assaying for substances which modulate utrophin promoter activity

Method 1:

This method uses a mouse *mdx*-H2K myoblast line stably transfected with a human 7.0kb utrophin promoter-luciferase construct. On day 1 myoblast cells transfected with the construct are plated out in 6-well dishes, with compound or DMSO-only for the negative controls.

4 x 6 well plates are used for every 3 compounds (the compounds are dissolved in DMSO and stored prior to use). For example, compound A, or B, or C were each added to 1 well, while the remaining 3 wells contain only DMSO. This results in 4 wells containing each compound and 12 wells with DMSO alone. Due to the inherent noise of both the harvesting/assay and cell seeding/growth steps, this is the minimum number that results in meaningful analysis. Setting up the plates in this way means that the data really are paired, and can be analysed with a paired student T test. This provides a more powerful statistical analysis rather than putting each compound on a different plate and comparing it with a control plate.

On Day 4 the cells are harvested and luciferase quantitation and pairwise analysis is carried out.

Method 2:

Compounds which up-regulate the endogenous utrophin promoter are be found using *mdx*-H2K myoblasts that are not transfected with the utrophin promoter-luciferase construct. *Mdx*-myoblasts can be used to mimic utrophin transcprition and protein stability in dystrophin-deficient cells.

Identification of utrophin protein expression

Quantitative Western Blotting is used to measure the level of

utrophin expression (Tinsley JM, *et al., Nature Medicine* **4**, 1441-1444.) Using 6 well plates and treating with compound as described above generates enough total protein sample to test by Western blotting. Antibodies specific to the A protein or B protein are used to quantify levels of either protein.

Identitification of utrophin RNA expression

Quantitative ribonuclease protection is used to analyse levels of utrophin expression. A pairwise design is used, as described above, but more cells are necessary. To see bands clearly, about 20-30μg total RNA is used. Each compound and control will need a 175 cm^2 tissue culture flask. A dual probe to simultaneously identify the A transcript and B transcript is be used.

Using the two techniques described compounds are identified after cell treatment which modulate utrophin levels. The same techniques are used for *in vivo* animal experiments where the compound is administered to dystrophin deficient *mdx* mice.

Interleukin-6 (IL-6) Interactions

Two related elements are present in the promoters of genes encoding acute phase proteins that mediate an increase in transcription stimulated by an IL-6 triggered signalling cascade (Hocke et al., 1992). One of these was found to be present in the exon 1B flanking sequence. Wild type and mutated reporter fusions for IL-6 were therefore tested for responsiveness in appropriate cell systems.

Constructs of the 1.5F B promoter normal and mutant (consensus change : ctggaa > gatatc[a] concerning the mutant: Hattori M et al (1990)*Proc. Natl. Acad. Sci. U S A.* Mar;87(6):2364-8.) were introduced into a promoter-less luciferase reporter vector and transfected into IN157 cells with a renilla firefly control. Cells were washed and charcoal stripped serum added 5 hours

post-transfection and left overnight. IL-6 amounts were added as illustrated with an appropriate amount of IL-6 soluble receptor. The cells were left for 24 hours and then assayed for activity using a luminometer.

A dosage dependent transcriptional response was noted in the normal, but not the mutated reporter construct (figure 10). This result indicates the existence of a cytokine mediated signalling pathway which causes up-regulation of the B utrophin promoter through the interaction of IL-6 and IL-6 receptor with the conserved IL-6 response element.

Discussion

We have demonstrated that there is a second promoter within intron 2 of the utrophin gene, driving expression of a unique first exon that splices into a common 13kb mRNA. These data are important, both in terms of understanding the molecular physiology of utrophin expression, and in view of their application to therapeutic intervention in DMD.

The functional consequences of genes having more than one promoter have been postulated (reviewed in (Ayoubi, *et al* (1996) *FASEB J.* **10**,453-460). A single gene may achieve a complex temporal and spatial expression pattern by interaction of different promoters with discrete subsets of transcription factors. Dystrophin is an example: three dissimilar promoters are active at different levels in specific cell types within the heart, skeletal muscle and the brain (Gorecki, *et al.* (1992) *Hum Mol Genet* **1**, 505-510., Barnea, *et al.* (1990) *Neuron* **5**, 881-888, Holder, *et al. Human Genetics* **97**, 232-239). Northern blot analysis, however, indicates that utrophin exon 1B is ubiquitously expressed, implying that promoters A and B are co-expressed in many tissues. It is conceivable that examination of transcript distribution in whole tissue samples has masked cell type-specific patterns of expression. Data

from isolated human cell lines *in vitro* support this notion; we observed large differences in promoter B activity between different cell lines, consistent with an *in vivo* expression profile involving specific cellular populations. Alternatively, the two promoters may be spatially regulated at a sub-cellular level. Within adult skeletal muscle fibres, promoter A is synaptically driven (Gramolini, *et al.* (1997) *J Biol Chem* **272**, 8117-20.), yet aggregates of utrophin mRNA are detectable at up to 25% extrasynaptic nuclei (Vater, *et al.* (1998) *Molecular and cellular Neuroscience* **10**, 229-242). Expression of promoter B in the extrasynaptic compartment might be invoked as one possible explanation.

A second proposed function of alternative promoters is the generation of transcripts with interchangeable 5' exons, giving rise to mRNAs with alternative 5'UTRs or proteins with novel N-terminal domains. Unlike exon 1B, utrophin exon 1A contains a long GC-rich 5'UTR. In some transcripts, GC-rich 5'UTRs are not translated efficiently (Kozak, M. (1991) *J Cell Biol* 115, 887-903.), and there are examples of genes in which alternative use of GC-rich and non-GC-rich 5'UTRs has been implicated in post-transcriptional regulation of protein synthesis (Nielson, et al. (1990) *J Biol Chem 265,* 13431-13434.). In addition, the predicted 31 amino acids encoded by exon 1B are different to the 26 amino acids of exon 2A; the functions of the resulting N-termini may be different.

The discovery of a second promoter provides a new target for the upregulation of utrophin to ameliorate the DMD phenotype. Promoter B is highly regulated, probably by different factors from promoter A, including IL-6. Elucidation of the mechanisms responsible for the large difference in promoter B activity between IN157 and HeLa cells might lead to identification of a factor that can be delivered to muscle to activate utrophin expression. Importantly, as the N-box motif is absent from promoter B, this is unlikely to carry any risk of NMJ

disruption potentially inherent in the pharmacological manipulation of synaptically regulated promoter A.